DOCUMENT RESUME

ED 279 183 FL 016 491

AUTHOR Tommola, Jorma
TITLE Simplified and Abridged Texts: What Happens to Content and Sentence Structure?
PUB DATE 84
NOTE 18p.; In: Ringbom, Hakan, Ed,; Rissanen, Matti, Ed. Proceedings from the Nordic Conference for English Studies (2nd, Hanasaari/Hanaholmen, Finland, May 19-21, 1983); see FL 016 498. Pages 5-6 contain small print.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; Content Analysis; *Difficulty Level; Discourse Analysis; Editing; *English; *Language Processing; Phrase Structure; *Reading Materials; *Sentence Structure; Statistical Analysis; *Structural Analysis (Linguistics)
IDENTIFIERS *Simplification (Language)

ABSTRACT

A comparison of an original and a simplified version of a text from the point of view of content and certain linguistic characteristics is examined. The text is a Daniel Defoe story. The effects of simplification on content are examined through analysis of propositions or "text atoms" of the text, their arrangement into an intermediate macrostructure, and the relation of this content structure with the content of the simplified version, especially to see how the simplified version followed the underlying content paths. The effects of simplification on sentence structure are measured by a number of mechanical counts (number of sentences, mean sentence and clause length, mean number of clauses per sentence, and the distribution of these numbers), sentence and clause types, depth of embedding, and subordination. Implications of simplification for language processing are also considered. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
N. Enkvist
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED279183

Jorma Tommola
Department of English
University of Turku

IMPLIFIED AND ABRIDGED TEXTS: WHAT HAPPENS TO CONTENT ND SENTENCE STRUCTURE?[1]

1. Introduction

The complexity or simplicity of text can be viewed as a combination of at least three factors: content (propositional structure), the language used to express that content (linguistic and stylistic properties of the surface text), and the listener or reader. Studying the first two factors yields descriptions of texts or accounts of differences between texts and text types. The function of psycholinguistic studies of text processing is to examine whether processing correlates can be found for the content or language parameters, i.e., whether these factors actually explain why some texts are simpler to process than others[2].

The explicit aim of producing a simplified version is to maximise processing ease. Such a text should thus reflect the re-writer's intuitions about, firstly, what elements of content are obligatory, and how propositional content should be organised to increase readability. Secondly, it will reflect the re-writer's intuitions about linguistic factors in the surface text that are likely to facilitate processing. Simplified and abridged texts are thus not without a certain interest, since their special characteristics can provide processing hypotheses to be studied experimentally.

In this paper, unsimplified and simplified text will be compared from the point of view of content and a small selection of linguistic characteristics. After a brief description of the material (section 2), some remarks are given on content structure (section 3) and sentence structure (section 4). Processing considerations are briefly touched upon in section 5.

FL016491

2. Material and method

The Life, Adventures and Piracies of Captain Singleton by Daniel Defoe[3] was chosen as the original (henceforth, O) text for three reasons: firstly, Defoe's sentence structure is interestingly complex, and the simplifier thus needs to introduce considerable changes; secondly, the text is a straightforward narrative in the first person with little dialogue; and thirdly, a convenient abridgement[3] (henceforth, the S version) exists in a series of graded readers.

In order to look at the development of content structure, a number of O episodes were examined for their content elements and functional relations between these, and similar analyses were carried out on the corresponding S episodes. One possible analysis of one pair of episodes is presented below.

To examine changes in sentence structure, samples of 2500 words were taken from the beginning of O and S. Though small, the samples seemed roughly adequate to bring out basic differences in some frequency distributions. The unit of observation was a clause, defined as in Quirk et al. 1972. The material contained 328 (O) and 394(S) clauses. The data was coded numerically using a 34-variable coding system, which concentrated on the subordination relations between clauses. Analyses were carried out using the BMDP statistical software package.

3. Content

What happens to content in simplification may be looked at in terms of the 'propositions' or 'text atoms' (cf. Enkvist 1975; van Dijk 1980) of the original text, their arrangement into an intermediate macrostructure, and the relation of this content structure with the content of the S version. The formalisation of propositions and the definition of their scope and the functions they realise is a major research topic in text studies (e.g. Mayer 1975; Frederiksen 1977; van Dijk 1979, 1980). For obvious reasons, a formal microlevel analysis (cf. Kintsch 1977) cannot and need not be attempted here.

For the present purpose, some episodes from O and S were taken apart into propositions, defined here as content units expressing a state of affairs, an action, or a thought, and corresponding fairly closely to the FACT representation of van Dijk (1980). Typically, the content of such a proposition tends to be expressed in a syntactic clause (cf. van Dijk 1979).

These propositions constitute a 'base list', an informal equivalent of the text base of Kintsch and van Dijk (1978). An intermediate macrostructural representation of the O content was constructed from the base list by creating a set of 'macropropositions' or summarising statements, for example by taking all propositions that enumerated the component parts of an action and constructing a more global statement from these. Van Dijk's macrorules (1977, 1980) represent an attempt to capture this process[4]. The associative functional relations between the resulting higher-level content units were determined, and the episode was organised into a temporal succession of macrostatements that carried the narrative forward. Other content units that had a subordinate or supplementary function were connected to these narrative points or nodes with links representing various types of relationship. The purpose of this enterprise was to attempt a description of the content network underlying the surface text of the O episode, in order to see how the simplifier traversed the paths of this network in the process of constructing the abridged version.

In the episode to be examined here (p. 20-23 in the O), the protagonist, Bob Singleton, together with other sailors, has been left on a tropical island after an unsuccessful mutiny, and now begins to consider ways of escaping. A total of 149 propositions were extracted; the following is a sample from the base list:

...
24. our consultations concerned the escape from this place
25. we were now upon the consultations
26. as I have said
27. (the consultations) ended in this (=30) only
28. we had two carpenters among us
29. they had tools of almost all sorts with them
30. we should try to build us a boat
31. we should go off to sea with (the boat)
32. then perhaps we might find our way back to Goa
33. we could land on some more proper place
...

Figure 1 (next two pages) gives one possible intermediate macrostructure of the O episode, and the arrangement of the content units chosen for the S version. Macropropositions are represented by numbers in angled brackets. Their content is also outlined in a few words. The numbers enclosed in curved brackets represent statements in the base list. A system of 12 associative functions[5] was used to describe the relations between the macrostructural content nodes.

FIGURE 1. The content structure of one episode in O and S.

(Fig. 1 continued)

The first thing to note is that while the O base list is 149 items long, the S base list contains 26 propositions. This indicates the extent of content reduction. Secondly, the S version is based on a small selection of NEXT-linked nodes. It is obvious that not all of the NEXT-linked nodes in the O have the same narrative or rhetorical height, and only the most central ones can be chosen for the S version. In a sense the S version thus represents a next higher version of content structure. The deletion of all but the central action nodes also results in the omission of most of the information that gives depth to the action sequence or the protagonist's character -- a feature which is also characteristic of stereotypical TV and comic-strip narratives. Similarly, META and ASIDE information is removed. For those action nodes that are selected, some supplementary information is present in most cases. Maximal reduction is thus not feasible, and the abridgement must still contain a sufficient amount of detail.

Simplifications are thus seen to resemble summaries or recall protocols (cf. van Dijk 1979) in that they contain some fairly high-level version of the content structure of the original. The base list propositions of the S version often tend to be summarising combinations of O propositions. Alternatively, and more commonly, it is those O propositions that directly express macrostatements that tend to get selected for the S base list. The base list propositions themselves are not essentially more complex in O as compared with those in S. The difference between O and S lies in the fact that the S representation is built up mainly of high-level generalised statements, and few supplementary propositions are linked with these.

The S version directly provides a higher-level content representation of the episode, and there is little danger that the reader would lose the thread of the story in a complex network of content units and associative relations. The comprehension of a narrative or any other type of text is, of course, a process of extracting and storing the central content structure. In the S version, this process is made easier for a reader who might otherwise need to direct most of his processing capacity to lower-order features of the text.

4. Sentence structure

Apart from the selection and organisation of content units, simplification also involves textualising these units in a certain way. Textualisation is largely independent of content structure; for example, one could easily take

the 26 items of the S base list and write them into a single sentence with excessively complex syntax. In this section, some observations will be given on what happens to sentence structure in simplification.

Sentence and clause length. The coding system used in the analysis of sentence structure included a number of mechanical counts. Some results from these are given in the following.

Firstly and obviously, simplification leads into a reduction in the number of sentences, in fact by about 50 per cent. Mean sentence length in words is reduced by a similar factor, from 60.2 (O) to 33.8 (S) words. Mean clause length is also reduced, from 7.35 words (O) to 6.36 words (S), but as can be seen, the reduction is fairly slight. Thus, as far as length in words is concerned, clauses appear to remain relatively constant. Basic clause elements must obviously be present, and clauses are reduced mainly by editing out optional adverbials or shortening coordinated NPs. The mean number of clauses per sentence is reduced from 8.20 (O) to 5.25 (S). The variable of clauses per sentence thus gets clearly different values in the two versions. It is interesting to note in this connection that Ellegård (1978) found no differences in this respect between the four text types he studied in the Brown corpus, which suggests that some existing differences may have been masked by sample size. Thus, a S version introduces some changes in the way clauses are combined, while the clause itself does not undergo extensive changes.

The frequency distributions of sentences and clauses of different lengths are given in Figures 2 and 3.

FIG. 2.

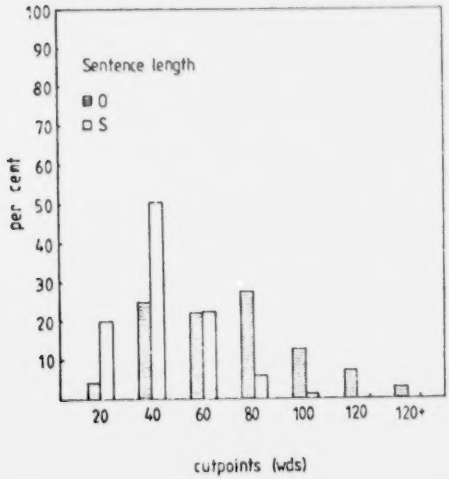
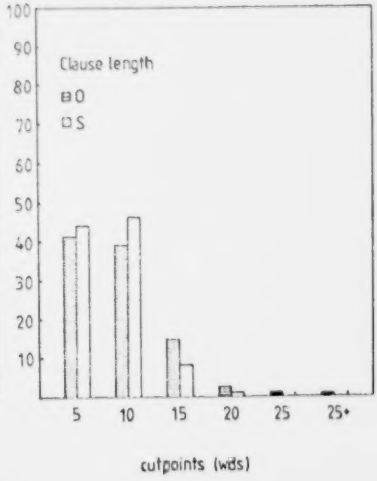


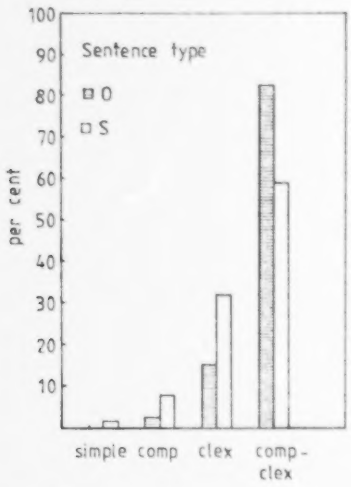
FIG. 3.



About 70 per cent of S sentences contain 40 words or less, whereas the O distribution is more even and also contains a number of sentences over 100 words long. The distribution of clause lengths changes relatively little, though there is a slight tendency for the S cases to concentrate even more clearly towards the shorter end, and though a small number of long clauses (with lists of coordinated object NPs, in fact) appears in the O version.

Sentence and clause types. Figure 4 gives the distribution of the sentences into the traditional sentence types for the two sets of material. The four types represent the pooled results from a variable which originally had six levels.

FIG. 4.



As can be seen, hardly any simple or compound sentences were present in the O sample, and the most frequent sentence type in O was the compound-complex sentence, often with complicated coordinations of subordinate clauses. Even in the simplified version, well over one half of the sentences have multiple main clauses with subordination. Compound-complex sentences are, however, clearly less frequent in S than in the O version, notably because there is a distinctly larger number of complex sentences with one main clause in S than in O. Simplification does not appear to involve the predominant use of simple sentences or coordinated main clauses. This would indeed be stylistically awkward, and would necessitate a considerable addition of redundancy, which would not contribute to clarity.

The distribution of clause types was studied in terms of the Quirk et al. (1972) categories. The variable of finite clause type contained 10 levels, including the usual classes of main clause and relative, nominal, adverbial and other dependent clauses. The variable of non-finite clauses was categorised into infinitival, participial, and verbless clauses, with various subtypes for the first two groups.

The percentages of the various types of finite and non-finite clause, calculated from the total number of clauses in S and O are set out below in Table 1.

TABLE 1. Percentages of clause types

	S		O	
FINITE CLAUSES:				
1st or only mcl	12.80		19.08	
2nd mcl	10.06		13.99	
3rd mcl	5.18	29.57	6.87	43.00
4th mcl	1.22		2.29	
5th mcl	0.30		0.76	
relative	17.07		11.45	
that-nominal	7.01		10.69	
wh-nominal	3.05	48.17	3.31	38.93
adverbial	16.46		12.47	
other	4.56		1.02	
NON-FINITE CLAUSES:				
infinitival	3.47		14.00	
ing-participle	6.71		3.82	
ed-participle	1.22	22.26	-	18.07
verbless	0.61		0.25	
	100.00		100.00	

Looking first at the types of finite and non-finite dependent clauses, it seems that moving from the O to the S version implies a reduction in the use of some clause types, for example relative and adverbial clauses, which typically contain modifying and elaborative information. On the whole, the reduction in the number of dependent clauses occurs chiefly for finite clauses. At the same time, there is an increase in the proportion of main clauses. Thus, as Figure 5 shows on the following page, simplification results in text where there are distinctly fewer subordinate clauses per main clause than in the original version. In S, information tends to be placed more often in main clauses, or alternatively, subordinate clauses with strictly speaking accessory information are edited out.

Figure 6 gives one more observation on clause types. This concerns the finiteness and non-finiteness of dependent clauses in the two samples. There

FIG. 5.

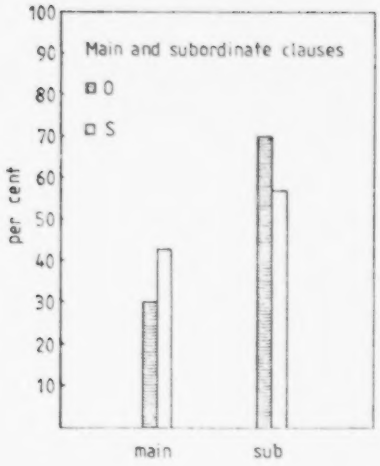
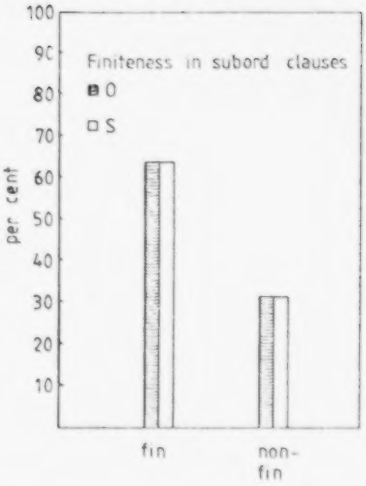


FIG. 6.



is no difference in the distributions, which suggests that redundancy addition by turning compact, non-finite constructions into finite structures either is not taking place, or, as actually seems to be the case, the O narrative contains few non-finite structures in need of simplification. This is probably connected with genre: the use of heavy participial modification is not likely to be as characteristic of narrative text as it might be of scientific expository writing, for example, where simplification might well demand the dilution of modifying structures by the manipulation of dependent clause finiteness.

Depth and subordination sets. The distribution of depth of embedding across all clauses is given in Figure 7. The pattern is reasonably clear: in the simplified version, the majority of clauses are either main clauses with depth 1 or embeddings at the shallower levels. For the O text, the cases are more evenly distributed across different depth levels. A crosstabulation of finiteness with depth indicates that in the original version it is particularly the non-finite clauses that are responsible for the more even distribution of depth; i.e., non-finite clauses often tend to occur deep down in the set of embeddings.

The embedded clauses suspended from a matrix clause may be divided into the usual four positional groups which will here be called subordination sets:

FIG. 7.

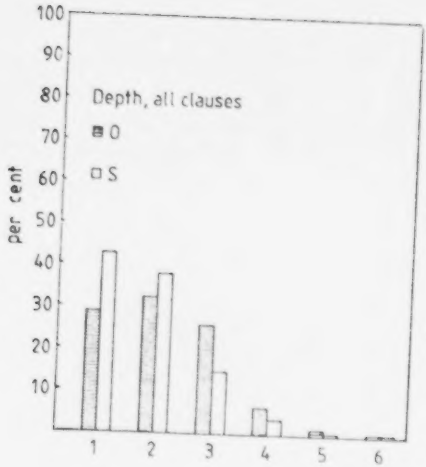
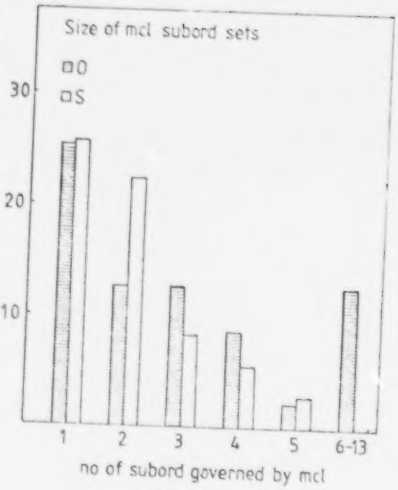


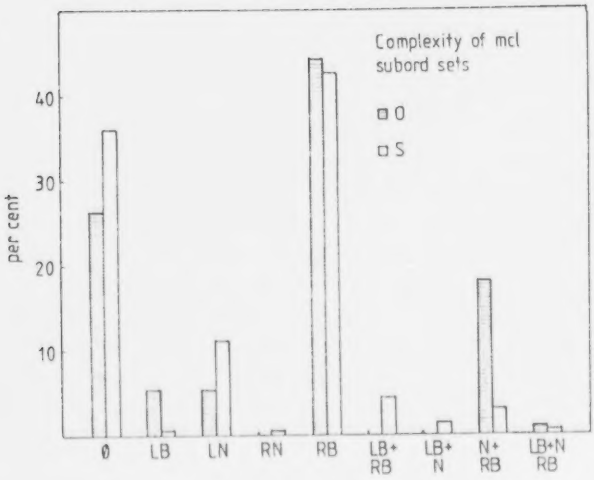
FIG. 8.



left-branching (LB), left-nested (LN), right-nested (RN), and right-branching (RB)[6]. One aspect of the complexity of the subordination sets is illustrated in Figure 8 for main clauses that function as matrix clauses. These are categorised into 6 groups according to the total number of embeddings suspended from them across all subordination set types. The considerable difference between O and S is perhaps best seen in the fact that there are several main clauses governing 6 to 13 dependent clauses in the original, and none in the S version.

The complexity of the main clause subordination sets is illustrated in terms of their types in Figure 9, where the percentages of nine such types are given. The first pair of columns (Ø) represents main clauses with no embeddings; the next four pairs of columns represent main clause cases where a single set of one of the four types is governed by the main clause; and the final four pairs represent combination cases. It may be observed from Figure 9 that there are more main clause cases without subordinates in S. Although the single LB embedding suspended from a main clause is infrequent on the whole, nearly all LB occurrences come from the O version. On the other hand, combination types involving LB embeddings appeared in S but not in O. The N+RB combination is clearly more frequent in O as compared with S. It is apparent for both the O and the S version that only a minority of

FIG. 9.



main clauses govern combinations of subordination sets. Thus, the sentence structure of the narrative text in the present case, though fairly complex, does not even approach the complexity of some other text types, especially law text. The obvious explanation lies in the demands of single-pass processing, which are briefly touched upon below.

4. Conclusion

In a simplified version, O propositions are collapsed into macropropositions or are deleted altogether. Thus a simplification resembles a precis or a recall protocol, where only the obligatory skeleton of the O content appears. This generalisation and reduction of semantic content has little effect on the surface clause, which actually tends to express a proposition in the present sense. But producing a simplified version immediately requires changes in the way propositional content is presented: fewer propositions tend to be expressed within the surface unit of the sentence. This results in the kinds of changes of sentence structure that were illustrated above: sentences become shorter though clauses do not, there are fewer clauses per sentence, the ratio of subordination changes, the overall depth of subordination becomes shallower, and the size, complexity, and combinatory

possibilities of subordination sets are reduced. The result is that the content units selected for S are offered to the reader in chunks designed not to overload the comprehension system.

The essence of fluent understanding in listening or reading is fast recoding of surface constituents into rich semantic units. In this sense, there is a definite time element even in reading: the surface string must be perceived, its constituents must be segmented, and the relations of the constituents must be determined without too much delay. If unanalysed, the input string vanishes from the working memory with frustrating speed, and regressions to previous parts of the text are needed to recover the lost information. This disrupts the construction of a gist representation. A skilled reader, who glances through the 80-word sentence corresponding to the base list fragment given above, has no problem in holding the propositional content and relations in his or her mind simultaneously. Readers for whom simplified versions are usually written, e.g. intermediate non-native users of the language, need to devote a considerable portion of their processing capacity to the handling of intra-clausal relations, vocabulary, morphological signals, and even letters. The handling of foreign-language structures requires conscious attention. The fact that the S version gives only the skeletal main points of a text, without including the complex relational network of supplementary propositions, makes it more likely that global aspects of text meaning can be processed even if the reader needs to give much of his processing capacity to local features.

Results reported by Perfetti and Lesgold (1977) are related with these considerations. In probe recall tests, differences between skilled and less skilled L1 comprehenders are negligible when test materials consist of one-clause sentences. Extensive differences are obtained when test sentences (after which the memory probe is presented) contain two clauses. Thus, the processing demands connected with the integration of clause content may be a source of individual differences in comprehension. With single-clause sentences, the less skilled comprehender can still handle even the beginning of the sentence, but encoding a second clause already requires more work, and both clauses become less available (Perfetti and Lesgold 1977:155). In the case of simplification, the reduction of clause combinations serves to reduce working memory load. Bhatia's (1983) 'easification' of legal study materials involves a related reorganisation of complex text into more manageable units, and also employs a diagrammatic presentation aimed at developing gist recognition strategies.

It is obvious that subject matter with a complicated propositional structure easily leads to complex syntax on parameters like the ones briefly mentioned above. But propositional complexity does not necessarily require the use of complicated syntax. Neither is it the case that precision in the expression of content units and their semantic relations would necessarily demand complex surface forms. The convention of, say, legislation, whereby all the necessary conditions and qualifications are included within the scope of one sentence is, however, understandable since compactness tends to be a virtue, and the resulting prose is not meant to be read in the same way as normal prose, i.e. at a single pass. Whenever single-pass processibility is the aim, the content needs to be organised and presented keeping in mind the inherent limitations of the working memory.

Accordingly it is useful to build up syntactic units using only a moderate number of propositions, and to organise the sentences into natural sequences governed by a clear overall text strategy. This increases the likelihood that the reader will be able to handle the content structure with ease. Making complicated texts simpler to process thus depends largely on how the content units are linearised on the surface. This is also noted by Shuy and Larkin (1978) in their discussion of insurance policy language. A similar idea underlies courses given to Finnish lawyers (cf. Huovinen-Nyberg 1981) where, among other things, court decisions are rewritten as texts rather than as sentences, often with a 'narrative' type of text strategy. The present material seemed to indicate that a similar principle underlies simplification and abridgement.

NOTES

1. This paper continues an earlier one (Tommola 1980), and is partly based on work done with students in 1979 in a seminar where several samples of narrative text were compared with published simplifications or abridgements. The main aim was to study changes occurring in sentence structure, notably in clausal embedding relations.

2. Various content and language parameters can be suggested as factors in simplicity or complexity. Traditional readability indices, for example, concentrated on word frequency and such mechanical variables as sentence length, the number of prepositional phrases per 100 words, etc. (Klare 1974 lists several such formulas). Later syntactic complexity formulas attempted to take into account linguistically more interesting factors derived from various versions of generative grammar. Cloze techniques approach readability from the point of view of the reader; their original motivation came from information theory, but the construct validity of cloze can also be derived

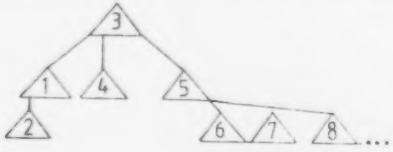
from considerations of the reader's pragmatic and linguistic expectancies. Recent experimental and other psycholinguistics has yielded an entire literature on linguistic factors which help or hinder the processing of text. Thus, for example, processing is affected by the need to make inferential bridges while constructing a mental representation of content (Clark and Haviland 1977), the semantic density of the material (Perfetti 1969, Kintsch 1977), or the extent to which paragraph structure helps the integration of new information with what was said before (e.g. Kieras 1978, Vande Kopple 1982). A syntactic factor that has received some attention is the number of future structural commitments that must be held in mind at a certain point, i.e., the question of certain forms of embedding as a source of syntactic complexity. This was represented in purely surface-structural terms by Yngve's depth measure; later research tends to consider this as a semantic question. Wanner and Maratsos (1978) address the same question in ATN terms and present experimental evidence for a HOLD hypothesis whereby transient memory load is increased in relative clauses by the necessity of holding the analysis of the matrix clause suspended while the embedded relative is being handled. According to a 'cue deletion' hypothesis, processing is also affected by the presence or absence of surface cues to underlying structure (e.g. Hakes 1972), which prevent the comprehender from being led down the garden path by canonical syntactic expectancies.

3. A 1951 edition, published by J.M. Dent & Sons, was used. The S version comes from a series called Easy Readers, published by Grafisk Forlag, Denmark, and a number of other publishers.

4. All the stages in such an analysis (delimitation of content elements, specification of the functional associative relations, and derivation of gist representations) tend to be subjective (Tommola and Neuvonen, forthcoming). Yet it should be possible, especially for a concrete narrative, to isolate the chief component actions with some reliability.

5. The NEXT relation is taken as a basically paratactic link which adds a new content unit at the next point in time. Additional associations often obtain betweeen nodes linked by NEXT, e.g. the RESULT relation between <21> and <23>. WHILE is used to refer to actions etc. taking place simultaneously. GOAL, REASON, and RESULT are self-explanatory; CONTRAST is an adversative and CONCESS a concessive relation. NOWTHAT is half-way between REASON and CONDITION (not needed here), and represents a function usually performed by clauses of circumstance (Quirk et al. 1972). SPECIF is a link connecting examples, specifications, or descriptive details to a superordinate node, while REPEAT denotes restatement of previous information. META and ASIDE are links which do not connect directly to content nodes; they merely signal the existence at some point of metatextual elements or other propositions detached from the actual content. For similar attempts to analyse the functional roles of text units of various sizes see e.g. Tirkkonen-Condit 1982 and Laurinen, forthcoming. Meyer 1975 and Lieber 1980 contain further discussion of functional roles.

6. The following example from the O version (shortened at the end) has a main clause (3) with a left-branching set (clauses 1 and 2), a right-branching set (clauses 5-), and a left-nested preverbal set consisting of clause 4.



1 Being now, (2 as I have said,) a considerable number of us and in condition to defend ourselves, 3 the first thing (4 we did) was 5 to give every one his hand 6 that we would not separate from one another upon any occasion whatsoever, 7 but that we would live and die together; 8 that ... (p. 18).

REFERENCES

Bhatia, V.K. 1983. "Simplification v. easification - the case of legal texts". Applied Linguistics 4. 42-54.

Clark, H.H. and Haviland, S.E. 1977. "Comprehension and the given-new contract". In Freedle, R.O. (ed.) Discourse Production and Comprehension. Norwood, N.J.: Ablex.

van Dijk, T.A. 1977. "Semantic macro-structures and knowledge frames in discourse comprehension". In Just, M.A. and Carpenter, P.A. (eds.) Cognitive Processes in Comprehension. Hillsdale, N.J.: John Wiley & Sons.

van Dijk, T.A. 1979. "Recalling and summarising complex discourse". In Burghardt, W. and Hölker, K. (eds.) Text Processing. Berlin: de Gruyter.

van Dijk, T.A. 1980. Macrostructures. Hillsdale, N.J.: Lawrence Erlbaum.

Ellegård, A. 1978. The Syntactic Structure of English Texts. Gothenburg Studies in English 43. University of Gothenburg.

Enkvist, N.E. 1975. Tekstilingvistiikan peruskäsitteitä. Helsinki: Gaudeamus.

Frederiksen, C.H. 1977. "Semantic processing units in understanding text". In Freedle, R.O. (ed.) Discourse Production and Comprehension. Norwood, N.J.: Ablex.

Hakes, D.T. 1972. "Effects of reducing complement constructions on sentence comprehension". Journal of Verbal Learning and Verbal Behavior 11. 278-286.

Huovinen-Nyberg, L. 1981. "Voiko helposti ymmärrettävän ja nopeasti luettavan tekstin kirjoittamista opettaa?" In Tommola, J. and Ruusuvuori, R. (eds.) Näkökulmia kirjoitettuun kieleen. Turku: AFinLA.

Vande Kopple, W.J. 1982. "The given-new strategy of comprehension and some natural expository paragraphs". Journal of Psycholinguistic Research 11. 501-520.

Kieras, D. 1978. "Good and bad structure in simple paragraphs: effects on apparent theme, reading time, and recall". Journal of Verbal Learning and Verbal Behavior 17. 13-28.

Kintsch, W. 1977. "Reading comprehension as a function of text structure". In Reber, A.S. and Scarborough, D.L. (eds.) Toward a Psychology of Reading. Hillsdale, N.J.: Lawrence Erlbaum.

Kintsch, W. and van Dijk, T.A. 1978. "Towards a model of discourse comprehension and production". Psychological Review 85. 363-394.

Klare, G. 1974. "Assessing readability". Reading Research Quarterly 10. 62-102.

Laurinen, L. (forthcoming). "Tekstin merkityssuhteiden analyysi". In Sajavaara, K. and Tommola, J. (eds.) Psykolingvistisiä kirjoituksia IV. Jyväskylä: AFinLA.

Lieber. P.E. 1980. Cohesion in ESL students' expository writing: a descriptive study. Diss. University of New York.

Meyer, B.F. 1975. The Organization of Prose and its Effects on Memory. Amsterdam: North-Holland.

Perfetti, C.A. 1969. "Lexical density and phrase structure depth as variables in sentence retention". Journal of Verbal Learning and Verbal Behavior 13. 70-79.

Perfetti, C.A. and Lesgold, A.M. 1977. "Discourse comprehension and sources of individual differences". In Just, M.A. and Carpenter, P.A. (eds.) Cognitive Processes in Comprehension. Hillsdale, N.J.: John Wiley & Sons.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. 1972. A Grammar of Contemporary English. London: Longman.

Shuy, R.W. and Larkin, D.L. 1978. "Linguistic considerations in the simplification and clarification of insurance policy language". Discourse Processes 1. 305-321.

Tirkkonen-Condit, S. 1982. "Makrotematiikka asiatekstin arviointikriteerinä". In Tommola, J. and Ruusuvuori, R. (eds.) Näkökulmia kirjoitettuun kieleen. Turku: AFinLA.

Tommola, J. 1980. "Some parameters of simplification". In Sajavaara, K., Räsänen, A., and Hirvonen, T. (eds.) Papers in Language Learning and Language Acquisition. Jyväskylä: AFinLA.

Tommola, J. and Neuvonen, P. (forthcoming)."Content structure and summaries of expository text in a foreign language". In Sajavaara, K. and Tommola, J. (eds.) Psykolingvistisiä kirjoituksia IV. Jyväskylä: AFinLA.

Wanner, E. and Maratsos, M. 1978. "An ATN approach to comprehension". In Halle, M,. Bresnan, J., and Miller, G.A. (eds.) Linguistic Theory and Psychological Reality. Cambridge, Mass.: The M.I.T. Press.